Joint Modelling of Event Counts and Survival Times

Jennifer Rogers
and
Jane Hutton and Karla Hemming
December 14, 2009

Abstract

We consider the analysis of data from the MRC Multicentre Trial for Early Epilepsy and Single Seizures (MESS), which was undertaken to assess the differences between two policies: immediate, or deferred treatment in patients with single seizures or early epilepsy. In studies of recurrent events, like epileptic seizures, there is typically lots of information about individuals' seizure patterns over a period of time, which is often not fully utilised in analysis. We develop methodology that allows pre-randomisation seizure counts and multiple post-randomisation survival times to be jointly modelled, assuming that both these outcomes are predicted by (unobserved) seizure rates. The joint model was found to be superior to standard survival methods, although interesting characteristics within the data, not present in the model were also highlighted. We consider modifications to the joint model to accommodate these properties.

Keywords: Epilepsy; Event rate; Recurrent events; Survival analysis

1 Introduction

The MRC Multicentre Trial for Early Epilepsy and Single Seizures (MESS) was undertaken to compare two treatment schemes: immediate versus delayed treatment in those patients presenting only a single, or few epileptic seizures. The effect of these policies on times to first and second seizures is assessed, whilst using available information we have on individuals' seizure history to enhance conclusions.

In studies of recurrent events, like epileptic seizures, there is frequently lots of information about individuals' seizure patterns over a period of time that is generally not fully utilised in analysis. Additionally, epilepsy is characterised by multiple seizures, not a single, isolated event, yet in many treatment studies it is often only time to first event that is analysed. We develop methodology that allows the pre-randomisation seizure counts and multiple post-randomisation survival times to be jointly modelled. This method assumes that all these outcomes are predicted by (unobserved) seizure rates, supposing that each patient has an underlying constant seizure rate, which we allow to vary depending on baseline attributes. Their subsequent post-randomisation seizure rate will be reduced relative to their associated baseline seizure rate, with a greater reduction resulting in a longer time

to seizure post-randomisation, indicating a better therapy.

The question of whether to start patients on a course of anticonvulsants after a single epileptic seizure remains an area of uncertainty (Chandra 1992). Antiepileptic drugs (AEDs) frequently come with unpleasant side effects which can include weight loss or weight gain, altered mood, drowsiness, hair loss, or even polycystic ovarian disease and teratogenicity. For most epilepsy sufferers the benefits of antiepileptic drugs will far outweigh the associated risks; however, for those individuals who have had only a single seizure, or have infrequent and mild epileptic seizures the question of whether to withhold treatment until absolutely necessary becomes an interesting one.

2 MESS

The MRC Multicentre Trial for Early Epilepsy and Single Seizures (MESS) addresses the question of immediate versus delayed treatment with antiepileptic drugs in those patients that have had one, or very few seizures. Interest lay in both the effects on short-term recurrence, and the long term prognosis.

MESS randomised 1443 individuals in the early stages of epilepsy to immediate or deferred treatment. The inclusion criterion was: being aged at least one month, having a suitably documented history of at least one clinically definite, unprovoked epileptic seizure and there being uncertainty in both clinician and patient as to whether treatment with AEDs should commence. Those allocated to the deferred treatment group had treatment withheld until both clinician and patient agreed that treatment was necessary. MESS was a pragmatic trial, meaning that all subsequent choices of antiepileptic drug, dose and duration were in line with the clinicians' usual practice. Baseline covariates collected at randomisation include the variables age, sex, and information surrounding patients' pre-randomisation seizure history, such as seizure type and seizure frequency. An electroencephalogram (EEG) was also requested for each individual. Detailed methods and primary analyses can be found in Marson et al. (2005) and Kim et al. (2006).

3 Joint Modelling of Event Counts and Survival Times

The MESS data arrives in the form of a pre-randomisation seizure count and times to first and second seizure post-randomisation. Most standard survival analysis may treat the pre-randomisation event count information as a covariate, possibly measuring this quantity as a covariate with error. As an alternative, Cowling et al. (2006) considered a technique that jointly models an individual's pre-randomisation seizure count, and a single post-randomisation failure time.

We describe how data in the form of an event count over a defined initial period, together with survival times, following a change in the event rate, can be modelled under a Poisson framework. The simplest model is a homogeneous Poisson process, which assumes that all individuals experience seizures according to a Poisson process with rate λ. Consequently the event count, X_i, for individual i, over a period u_i will be Poisson with mean λu_i, and the interevent times will be Exponential with rate λ. Count data are often overdispersed, rendering this model unsuitable. Some of the overdispersion may be attributed to individuals' covariates, so we allow the rate to vary with the covariates. Therefore, we let the rate for individual i be λ_i, which relates in some way to the individuals' covariates.

There may be additional overdispersion caused by natural heterogeneity in the population, often observed in epilepsy data. Cowling et al. (2006) discusses the use of the Negative Binomial distribution for count data. Suppose that we assume a Poisson process for seizures, with rate $\lambda_i\nu_i$, where λ_i is related to the individuals baseline covariates, and ν_i is a random term that follows a Gamma distribution with expectation 1, and variance $1/\alpha$. The parameter α measures the degree of heterogeneity in the population, with smaller values indicating higher levels of heterogeneity. Consequently, the event count over a period u_i for individual i, X_i, follows a Poisson distribution with rate $\lambda_i\nu_i$. Interevent times are then Exponential with the same rate, and due to the memoryless property of this distribution, the post-randomisation survival times are also Exponential, with the rate updated to allow for a treatment effect. It is assumed that the treatment effect acts multiplicatively on the rate, so that the post-randomisation interevent times for individual i are Exponential with event rate $\lambda_i\psi_i\nu_i$, where ψ_i depends on the treatment in some way.

A natural extension to the model proposed by Cowling et al. (2006) considers two seizures post-randomisation together with the pre-randomisation event count. Let T_{1i} and T_{2i} be the times to first and second seizure respectively and set $Y_{1i} = T_{1i}$ and $Y_{2i} = T_{2i} - T_{1i}$, so that Y_{1i} is the time to first seizure and Y_{2i} is the time from first seizure to the second. Both Y_{1i} and Y_{2i} will be independent and Exponentially distributed with rate $\lambda_i\psi_i\nu_i$. In summary:

$$\begin{aligned} X_i \mid \nu_i &\sim \text{Poisson}(\lambda_i u_i \nu_i), \\ Y_{ji} \mid \nu_i &\sim \text{Exponential}(\lambda_i \psi_i \nu_i), \qquad j = 1, 2, \\ \nu_i &\sim \text{Gamma}(\alpha, \alpha). \end{aligned}$$

The joint density of the survival times is the product of the densities of Y_{1i} and Y_{2i}, so that the joint model is specified by the following equations:

$$\begin{aligned} f_{X|\nu}(x_i \mid \nu_i; \lambda_i, u_i) &= \frac{(\lambda_i u_i \nu_i)^{x_i} \exp(-\lambda_i u_i \nu_i)}{x_i!}, \\ f_{Y_1,Y_2|\nu}(y_{1i}, y_{2i} \mid \nu_i; \lambda_i, \psi_i) &= (\lambda_i \psi_i \nu_i)^2 \exp(-\lambda_i \psi_i \nu_i (y_{1i} + y_{2i})), \end{aligned}$$

$$g_\nu(\nu_i;\alpha) \;=\; \frac{\alpha^\alpha \nu_i^{\alpha-1}\exp(-\alpha\nu_i)}{\Gamma(\alpha)},$$

where $\lambda_i = \exp(\boldsymbol{\beta}_1'\boldsymbol{z}_{1i})$ and $\psi_i = \exp(\boldsymbol{\beta}_2'\boldsymbol{z}_{2i})$.

If the random effect term is integrated out of the joint density of X_i and ν_i, the resulting unconditional density, $f_X(x_i;\lambda_i,u_i,\alpha)$, is simply the Negative Binomial. The unconditional joint distribution of the Y_{ji}, $j=1,2$, obtained when the random effect term is integrated out of the joint density of Y_{1i}, Y_{2i} and ν_i, turns out to be the bivariate Lomax distribution (Nayak 1987), with the following density and survivor functions:

$$\begin{aligned} f_{Y_1,Y_2}(y_{1i},y_{2i};\lambda_i,\psi_i,\alpha) &= \int_0^\infty f_{Y_1,Y_2|\nu}(y_{1i},y_{2i}\mid\nu_i;\lambda_i,\psi_i)g_\nu(\nu_i;\alpha)d\nu_i \\ &= \frac{\alpha+1}{\alpha}(\lambda_i\psi_i)^2\left\{1+\frac{\lambda_i\psi_i(y_{1i}+y_{2i})}{\alpha}\right\}^{-(\alpha+2)}, \qquad (1) \\ S_{Y_1,Y_2}(y_{1i},y_{2i};\lambda_i,\psi_i,\alpha) &= \int_{y_{2i}}^\infty\int_{y_{1i}}^\infty f_{Y_1,Y_2}(u,v;\lambda_i,\psi_i,\alpha)du\,dv \\ &= \left\{1+\frac{\lambda_i\psi_i(y_{1i}+y_{2i})}{\alpha}\right\}^{-\alpha}. \qquad (2) \end{aligned}$$

Each of the Y_{ji} also have univariate Lomax marginal distributions, each with density:

$$f_{Y_j}(y_{ji};\lambda_i,\psi_i,\alpha) = \frac{\lambda_i\psi_i}{(1+\lambda_i\psi_i y_{ji}/\alpha)^{\alpha+1}}, \qquad j=1,2. \qquad (3)$$

When formulating the likelihood we need to consider the different ways that censoring can occur. There are three different ways censoring can arise in this setting, namely: (i) Y_{1i} and Y_{2i} are both observed, (ii) Y_{1i} is observed, but Y_{2i} is censored, and (iii) Y_{1i} is censored, so Y_{2i} is taken to be censored at zero. We now consider these three situations separately:

(i) Y_{1i} and Y_{2i} both observed

In this situation the joint density of Y_{1i} and Y_{2i} contributes towards the likelihood, giving

$$\begin{aligned} &\int_0^\infty f_{X|\nu}(x_i\mid\nu_i;\lambda_i,u_i)f_{Y_1,Y_2|\nu}(y_{1i},y_{2i}\mid\nu_i;\lambda_i,\psi_i)g_\nu(\nu_i;\alpha)d\nu_i \\ &\qquad = \frac{(\lambda_i u_i)^{x_i}}{x_i!}\frac{(\lambda_i\psi_i)^2\alpha^\alpha}{\Gamma(\alpha)}\frac{\Gamma(x_i+\alpha+2)}{(\lambda_i u_i+\lambda_i\psi_i(y_{1i}+y_{2i})+\alpha)^{x_i+\alpha+2}}. \qquad (4) \end{aligned}$$

(ii) Y_{1i} observed, but Y_{2i} censored

In this situation the density of Y_{1i} and the survivor function for Y_{2i} contribute to the likelihood, giving

$$\int_0^\infty f_{X|\nu}(x_i \mid \nu_i; \lambda_i, u_i) f_{Y_1|\nu}(y_{1i} \mid \nu_i; \lambda_i, \psi_i) S_{Y_2|\nu}(y_{2i} \mid \nu_i; \lambda_i, \psi_i) g_\nu(\nu_i; \alpha) d\nu_i$$

$$= \frac{(\lambda_i u_i)^{x_i}}{x_i!} \frac{\lambda_i \psi_i \alpha^\alpha}{\Gamma(\alpha)} \frac{\Gamma(x_i + \alpha + 1)}{(\lambda_i u_i + \lambda_i \psi_i (y_{1i} + y_{2i}) + \alpha)^{x_i + \alpha + 1}}. \tag{5}$$

(iii) Y_{1i} censored, so Y_{2i} taken to be censored at zero

In this situation the survivor functions of Y_{1i} and Y_{2i} will contribute to the likelihood, however we assume that the second survival time is censored at zero, giving $S_{Y_2|\nu}(0 \mid \nu_i; \lambda_i, \psi_i) = 1$.

$$\int_0^\infty f_{X|\nu}(x_i \mid \nu_i; \lambda_i, u_i) S_{Y_1|\nu}(y_{1i} \mid \nu_i; \lambda_i, \psi_i) g_\nu(\nu_i; \alpha) d\nu_i$$

$$= \frac{(\lambda_i u_i)^{x_i}}{x_i!} \frac{\alpha^\alpha}{\Gamma(\alpha)} \frac{\Gamma(x_i + \alpha)}{(\lambda_i u_i + \lambda_i \psi_i y_{1i} + \alpha)^{x_i + \alpha}}. \tag{6}$$

Conversely, by keeping $S_{Y_2|\nu}(y_{2i} \mid \nu_i; \lambda_i, \psi_i)$ in the calculations we subsequently obtain a simpler likelihood function, so we proceed in this way:

$$\int_0^\infty f_{X|\nu}(x_i \mid \nu_i; \lambda_i, u_i) S_{Y_1|\nu}(y_{1i} \mid \nu_i; \lambda_i, \psi_i) S_{Y_2|\nu}(y_{2i} \mid \nu_i; \lambda_i, \psi_i) g_\nu(\nu_i; \alpha) d\nu_i$$

$$= \frac{(\lambda_i u_i)^{x_i}}{x_i!} \frac{\alpha^\alpha}{\Gamma(\alpha)} \frac{\Gamma(x_i + \alpha)}{(\lambda_i u_i + \lambda_i \psi_i (y_{1i} + y_{2i}) + \alpha)^{x_i + \alpha}}. \tag{7}$$

Note that equations (6) and (7) are equivalent when $y_{2i} = 0$.

Now let δ_{ji} be the indicator function for the jth survival time, taking the value 1 if the seizure is observed, and zero if the survival time is censored. Combining these indicator functions with equations (4), (5) and (7) allows the formulation of the log-likelihood for the observed data $\mathcal{D}$ on all the n individuals, given by

$$\begin{aligned} \ell(\alpha, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2 \mid \mathcal{D}) \;=\; & \sum_{i=1}^{n} \Bigg\{ \left[\sum_{k=0}^{x_i - 1} \ln(\alpha + k) \right] + (x_i + \delta_{1i}(1 + \delta_{2i})) \ln(\lambda_i) + x_i \ln(u_i) - \ln(x_i!) \\ & + \alpha \ln(\alpha) + \delta_{1i}(1 + \delta_{2i}) \ln(\psi_i) + \delta_{1i} \ln(x_i + \alpha) + \delta_{1i}\delta_{2i} \ln(x_i + \alpha + 1) \\ & - (x_i + \alpha + \delta_{1i}(1 + \delta_{2i})) \ln\left(\lambda_i u_i + \lambda_i \psi_i (y_{1i} + y_{2i}) + \alpha\right) \Bigg\}. \end{aligned} \tag{8}$$

It is straightforward to obtain the first and second derivatives of this function, allowing inference on the parameters α, $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$, using a numerical method such as Newton-Raphson.

4 Exploratory Analysis

Exploratory analysis was carried out on 1425 individuals; 18 were removed due to missing information, assumed missing completely at random. A further five individuals with incomplete information on their pre-randomisation seizure history were excluded from the statistical modelling. Of the 1425 individuals included in the exploratory analysis, 691 (48.91%) experienced at least one seizure following randomisation, with a subsequent 480 (69.46%) of these experiencing a second. We consider Kaplan-Meier plots of the outcomes time to first seizure and time from first seizure to second, examining possible treatment policy, EEG outcome and seizure type effects. The pre-randomisation seizure types are categorised as follows:

Tonic-Clonic - those presenting tonic-clonic seizures only pre-randomisation,

2° **Tonic-Clonic** - those presenting partial seizures followed by secondary tonic-clonic seizures pre-randomisation,

Generalised - those presenting any types of generalised seizures pre-randomisation (this group could include those having a combination of tonic-clonic and other generalised seizures),

Partial - those presenting partial seizures only pre-randomisation (either simple or complex),

Other - those presenting seizures pre-randomisation that do not fit into any of the above categories.

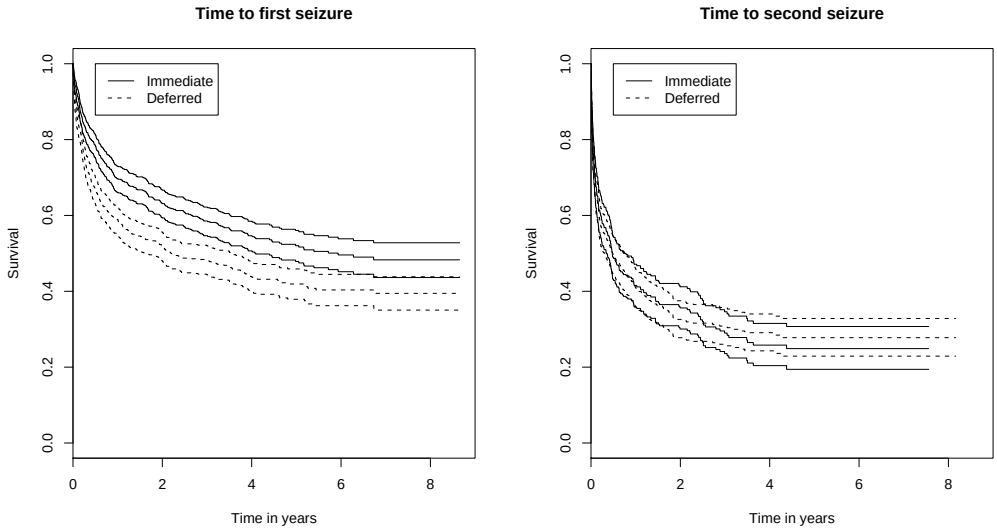

Figure 1: Kaplan-Meier curves for time to first seizure and time from first to second seizure (with 95% CI), stratified by treatment policy.

The Kaplan-Meier curves in Figure 1 highlight immediately that treatment policy appears to be influential in determining an individual's time to first seizure post-randomisation, but not their time from first to second seizure. A plausible explanation for this is that those individuals randomised to deferred treatment who experience a seizure post-randomisation would most likely receive subsequent treatment with AEDs, bringing them in line with those allocated to immediate treatment thereafter.

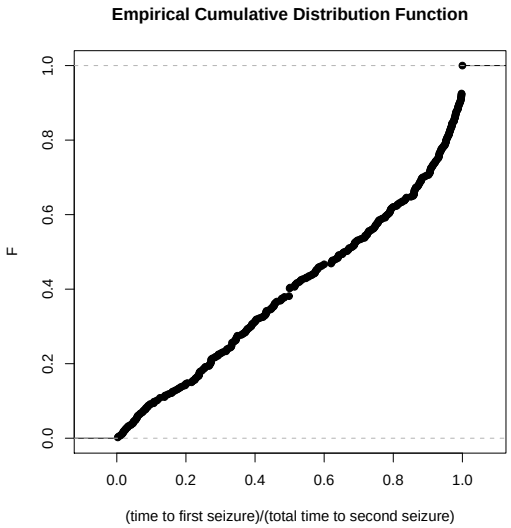

Figure 2: Empirical cdf for Y_1/T_2.

Figure 2 shows the empirical cdf for Y_1/T_2. Note that this has its median at 0.663, which suggests that for those experiencing at least two seizures post-randomisation, their time from first seizure to second is typically shorter than their time from randomisation to first seizure. Furthermore, around 60% of those having at least two seizures post-randomisation have $Y_1 > Y_2$, with approximately having Y_1 around nine times bigger than Y_2. These results suggest that clustering within seizures may be evident.

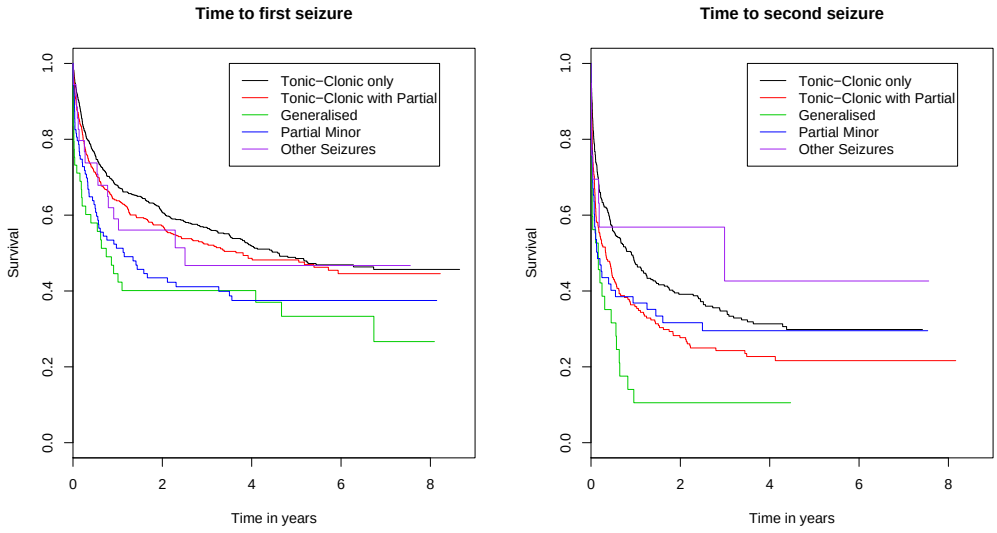

Figure 3: Kaplan-Meier curves for time to first seizure and time from first to second seizure, stratified by seizure type pre-randomisation.

Figure 3 shows that those experiencing generalised or partial seizures pre-randomisation typically have a shorter time to first seizure post-randomisation than the other seizure types. Individuals with generalised seizures pre-randomisation also present their second seizure post-randomisation much sooner than other seizure types. Additionally, the differences between the Kaplan-Meier curves appear to be more pronounced for the second seizure post-randomisation, than for time to first seizure.

Figure 4 suggests that for those participants with partial seizures only pre-randomisation, treatment policy appears to have no effect on their time to first seizure post-randomisation. For all other seizure types immediate treatment is favoured.

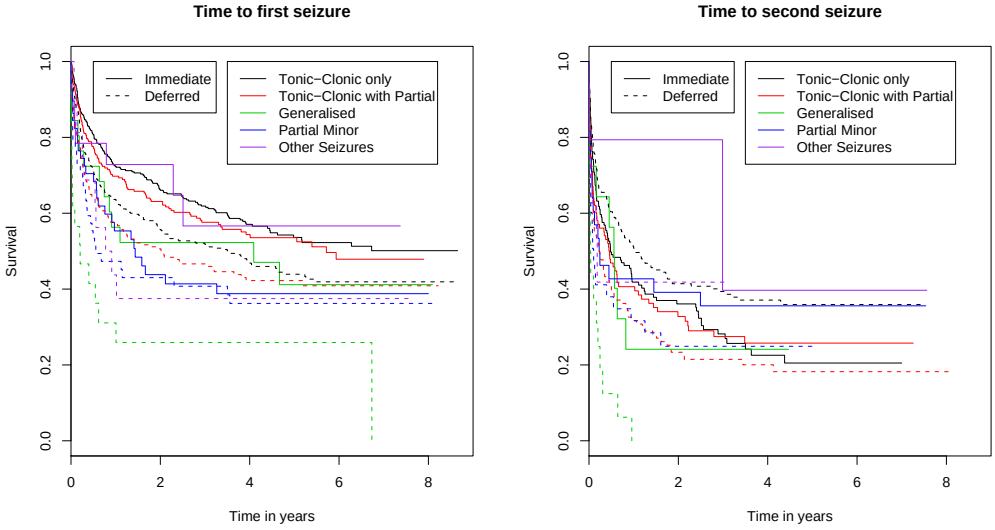

Figure 4: Kaplan-Meier curves for time to first seizure and time from first to second seizure, stratified by seizure type pre-randomisation and treatment policy.

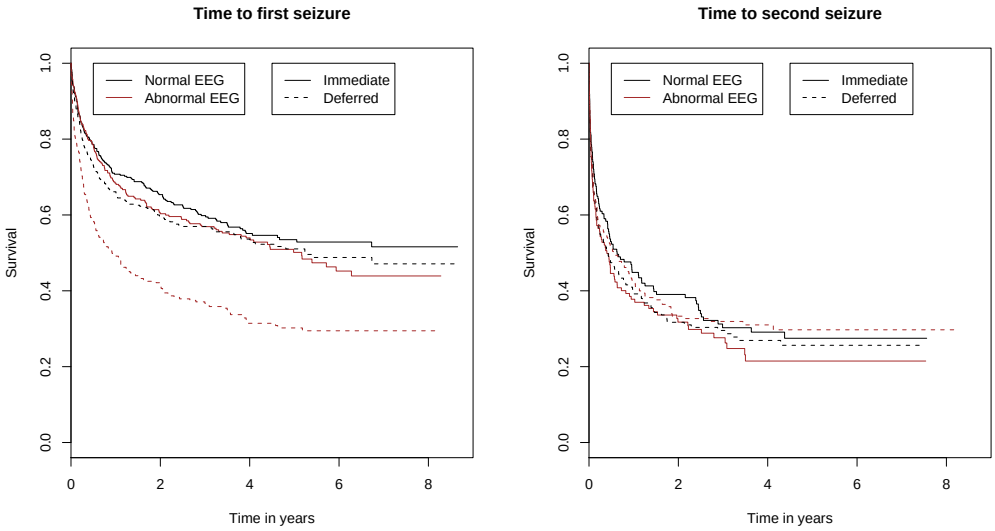

Figure 5: Kaplan-Meier curves for time to first seizure and time from first to second seizure, stratified by EEG outcome and treatment policy.

When considering EEG outcome, Figure 5 indicates that for those presenting a normal

EEG, treatment policy has no effect on their associated time to first seizure. For those with an abnormal EEG, allocation to immediate treatment brings their expected time to first seizure in line with those having a normal EEG. Those randomised to deferred treatment, following an abnormal EEG outcome can expect a much worse outcome. For time from first to second seizure post-randomisation there appears to be no difference in the four Kaplan-Meier curves.

Finally, Figure 6 gives an indication as to any interactions between EEG outcome and pre-randomisation seizure types that may be present. For time to first seizure, EEG outcome appears to be influential for those with secondary tonic-clonic seizures pre-randomisation, with those having a normal EEG faring better. EEG outcome also seems to have a slight impact on time to first seizure for those with tonic-clonic seizures only pre-randomisation, and possibly for those with generalised seizures pre-randomisation. These interactions are not seen for time from first to second seizure.

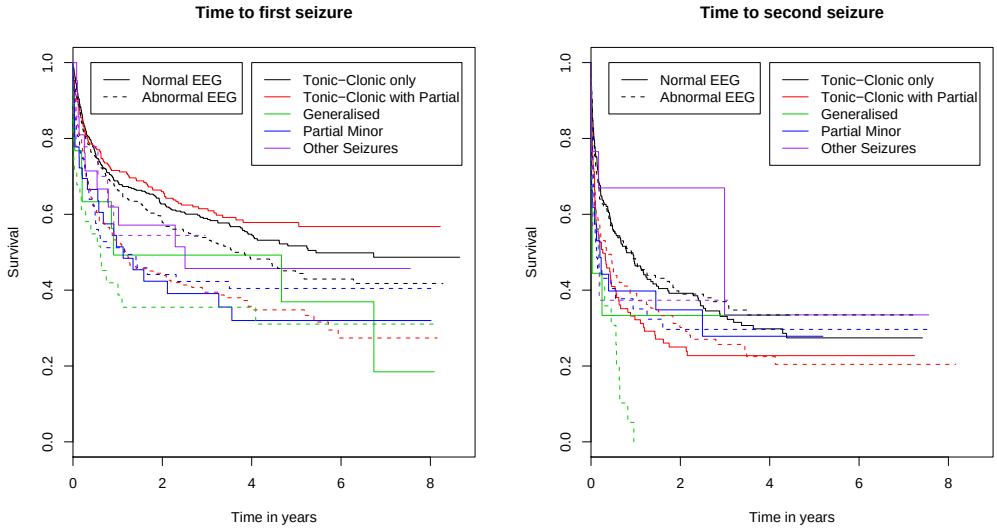

Figure 6: Kaplan-Meier curves for time to first seizure and time from first to second seizure, stratified by seizure type pre-randomisation and EEG outcome.

5 Results

The maximum likelihood parameter estimates obtained after considering marginal models for the count data and survival data separately are presented, followed by the implementation of the joint model, developed in §3. The estimated regression coefficients are also used to calculate estimates of the pre-randomisation seizure rate, $\widehat{\lambda}_i$, and seizure rate modifier, $\widehat{\psi}_i$, for the different treatment policies, EEG outcomes and seizure types pre-randomisation. Neither sex or age were found to be significant in any of the fitted models, so we exclude these variables completely.

5.1 Analysis of Pre-Randomisation Counts

The Negative Binomial GLM is considered as a marginal count model, specified by the following equation:

$$f_X(x_i; \lambda_i, u_i, \alpha) = \frac{\Gamma(x_i + \alpha)}{x_i! \Gamma(\alpha)} \left(\frac{\lambda_i u_i}{\alpha + \lambda_i u_i} \right)^{x_i} \left(\frac{\alpha}{\alpha + \lambda_i u_i} \right)^{\alpha}, \tag{9}$$

where $\lambda_i = \exp(\boldsymbol{\beta}_1' \boldsymbol{z}_{1i})$. Here $\boldsymbol{z}_{1i}$ is a vector of covariates for individual i, and $\boldsymbol{\beta}_1$ is a vector of regression coefficients, including an intercept term.

The estimated regression coefficients for the Negative Binomial marginal count model are given in Table 1. The small value of α suggests that there is substantial heterogeneity within the population. We can also see that those experiencing other seizures pre-randomisation should not expect to have a pre-randomisation seizure rate significantly different from those presenting partial seizures only.

Regression Coefficient	Estimates (standard errors) for Negative Binomial GLM
α	0.696 (0.027)
$\beta_{1,0}$	-2.630 (0.130)
$\beta_{1,t-c}$	-0.564 (0.141)
$\beta_{1,2^\circ t-c}$	-0.527 (0.147)
$\beta_{1,gen}$	0.611 (0.225)
$\beta_{1,par}$	reference
$\beta_{1,oth}$	-0.370 (0.258)
$-$Log-likelihood (d.f.)	3687 (1414)

Table 1: Estimated regression coefficients for the Negative Binomial GLM. A regression coefficient > 0 (< 0) would indicate an increased (decreased) seizure rate relative to the seizure rate in the reference group.

The regression coefficients also suggest that those having tonic-clonic, or secondary tonic-clonic seizures can expect to have similar seizure rates pre-randomisation. Additionally, under this model, those presenting generalised seizures should expect to have the highest pre-randomisation seizure rate.

5.2 Standard Survival Distributions

Transformations of the Kaplan-Meier estimates of the survivor function suggested that the survival data may be well modelled through a distribution that belongs to the accelerated failure time family of distributions, with further investigation supporting the use of the Log-logistic distribution. Additionally, recall that in the joint model, developed in §3, the unconditional distribution of each of the Y_{ij} turned out to be the Lomax distribution.

We therefore fit these two survival distributions to the post-randomisation survival times, namely time to first seizure and time from first to second seizure. The Log-logistic and Lomax distributions are defined by the following equations:

- Log-logistic (shape=γ, scale=$1/\mu_i$)

$$f_Y(y_i;\mu_i,\gamma) = \frac{\mu_i\gamma(\mu_i y_i)^{\gamma-1}}{(1+(\mu_i y_i)^{\gamma})^2}, \tag{10}$$

- Lomax (shape=γ, scale=γ/μ_i)

$$f_Y(y_i;\mu_i,\gamma) = \mu_i\Big(\frac{\gamma}{\gamma+\mu_i y_i}\Big)^{\gamma+1}, \tag{11}$$

where in each model $\mu_i = \exp(\boldsymbol{\theta}'\boldsymbol{w}_i)$, for a vector $\boldsymbol{\theta}$ of regression coefficients, including an intercept term, and a vector $\boldsymbol{w}_i$ of covariates for each individual i. Increasing values of the m regression coefficients, θ_k, $k = 0,\ldots,m$, correspond to an increase in the acceleration factor, and hence a decrease in the expected time to seizure. Conversely, negative values correspond to deceleration and a subsequent increase in the expected time to seizure. The parameter $\gamma > 0$ is a shape parameter and represents the degree of additional heterogeneity within the population, with smaller values indicating higher levels of heterogeneity. Recall that exploratory analysis supported the use of the Log-logistic distribution and note that equations (10) and (11) are equivalent when $\gamma = 1$. An estimate of γ close to 1 (indicating that there is considerable heterogeneity in the population) would therefore suggest that the data could be sufficiently modelled through the Lomax distribution, validating the use of the model developed in §3. We now present the estimated regression coefficients obtained when each of these distributions is fit to the post-randomisation survival times, time to first seizure and time from first to second seizure.

5.2.1 Time to First Seizure

The parameter estimates for time to first seizure only, for the Log-Logistic and Lomax distributions are given in Table 2.

Statistical significance of the covariates was determined by Wald tests (Wald 1943). In general, the Wald test is conducted by computing the Wald test statistic, which is given by $\widehat{\theta}_k^2/\mathrm{var}(\widehat{\theta}_k)$, for each of the $k = 0,\ldots,m$ estimated regression coefficients. Statistical significance is then determined by comparing this value against the chi-squared distribution. Wald tests on each of the estimated regression coefficients given in Table 2 show that for both survival distributions the only significant covariates are θ_{t-c}, $\theta_{2^\circ t-c}$ and $\theta_{eeg\times trt}$. Furthermore, those experiencing tonic-clonic only and secondary tonic-clonic seizures pre-randomisation can typically expect to have a longer time to first seizure post-randomisation. Additionally $\theta_{2^\circ t-c\times eeg}$ is significant in the Log-logistic model only. The lack of significant

covariates in these models is contrary to the observations made in §4. Probably most surprising is that the exploratory analysis suggested that treatment policy should be significant, but this is not being seen in the coefficient estimates.

Regression Coefficient	Estimates (standard errors) for the following models:	
	Log-logistic	Lomax
θ_0	-5.967 (0.645)	-3.785 (0.649)
θ_{trt}	0.438 (0.749)	0.216 (0.807)
θ_{t-c}	-1.255 (0.624)	-1.438 (0.623)
$\theta_{2^\circ t-c}$	-1.337 (0.647)	-1.307 (0.653)
θ_{gen}	0.758 (1.053)	0.650 (1.119)
θ_{par}	reference	reference
θ_{oth}	-0.237 (1.052)	-1.330 (0.963)
θ_{eeg}	0.364 (0.667)	0.252 (0.715)
$\theta_{\ln(rate)}$	0.059 (0.066)	0.145 (0.074)
$\theta_{t-c\times trt}$	-0.540 (0.667)	-0.535 (0.716)
$\theta_{2^\circ t-c\times trt}$	-0.717 (0.692)	-0.550 (0.741)
$\theta_{gen\times trt}$	-2.104 (1.116)	-1.283 (1.329)
$\theta_{par\times trt}$	reference	reference
$\theta_{oth\times trt}$	-1.346 (1.275)	0.182 (1.363)
$\theta_{eeg\times trt}$	-1.186 (0.350)	-1.087 (0.363)
$\theta_{\ln(rate)\times trt}$	0.037 (0.096)	0.008 (0.103)
$\theta_{t-c\times eeg}$	0.555 (0.687)	0.484 (0.739)
$\theta_{2^\circ t-c\times eeg}$	1.765 (0.714)	1.148 (0.768)
$\theta_{gen\times eeg}$	1.511 (1.172)	1.537 (1.331)
$\theta_{par\times eeg}$	reference	reference
$\theta_{oth\times eeg}$	0.332 (1.321)	1.627 (1.515)
γ	0.592 (0.019)	0.211 (0.007)
$-$Log-likelihood (d.f.)	5480 (1401)	5463 (1401)

Table 2: Estimated regression coefficients for survival models fitted to the times to first seizure. A regression coefficient, θ_k, > 0 (< 0) would indicate an increase (decrease) in the acceleration factor (hence a decrease (increase) in the time to event) relative to the seizure rate in the reference group. The reference group contains individuals with partial seizures pre-randomisation, with a normal EEG and randomised to deferred treatment.

The Log-logistic and Lomax distributions can not be compared using the standard likelihood ratio test, as they are non-nested models. The Akaike Information Criterion (AIC) (Sahamotoa et al. 1986) is a method for comparing two non-nested models and is given by $2(m - \ell)$, where m is the number of parameters in the model and ℓ is the maximised log-likelihood associated with the model. Computing the AIC for each of the two survival

distributions we have that for time to first seizure the AIC for the Log-logistic model is 10998, and for the Lomax distribution the corresponding AIC is 10964. This indicates that the Lomax distribution preferred over the Log-logistic distribution.

5.2.2 Time from First to Second Seizure

Regression Coefficient	Estimates (standard errors) for the following models:	
	Log-logistic	Lomax
θ_0	-3.319 (0.899)	-1.678 (0.765)
θ_{trt}	-1.326 (1.047)	-0.813 (0.972)
θ_{t-c}	-2.456 (0.865)	-2.595 (0.743)
$\theta_{2^\circ t-c}$	-0.852 (0.893)	-1.340 (0.774)
θ_{gen}	0.057 (1.378)	-0.535 (1.260)
θ_{par}	reference	reference
θ_{oth}	-2.249 (1.734)	-1.917 (1.355)
θ_{eeg}	-0.885 (0.899)	-1.126 (0.813)
$\theta_{\ln(rate)}$	0.045 (0.095)	0.060 (0.091)
$\theta_{t-c\times trt}$	2.024 (0.908)	1.323 (0.840)
$\theta_{2^\circ t-c\times trt}$	0.870 (0.935)	0.700 (0.865)
$\theta_{gen\times trt}$	-1.141 (1.388)	-1.284 (1.382)
$\theta_{par\times trt}$	reference	reference
$\theta_{oth\times trt}$	-0.517 (2.196)	-1.645 (1.883)
$\theta_{eeg\times trt}$	0.775 (0.487)	0.574 (0.468)
$\theta_{\ln(rate)\times trt}$	0.117 (0.141)	0.116 (0.136)
$\theta_{t-c\times eeg}$	0.456 (0.929)	0.792 (0.855)
$\theta_{2^\circ t-c\times eeg}$	0.195 (0.961)	0.617 (0.883)
$\theta_{gen\times eeg}$	1.204 (1.505)	0.978 (1.477)
$\theta_{par\times eeg}$	reference	reference
$\theta_{oth\times eeg}$	1.939 (2.213)	2.489 (1.968)
γ	0.595 (0.024)	0.274 (0.012)
$-$Log-likelihood (d.f.)	2993 (1401)	2973 (1401)

Table 3: Estimated regression coefficients for survival models fitted to the times from first to second seizure. A regression coefficient, θ_k, > 0 (< 0) would indicate an increase (decrease) in the acceleration factor (hence a decrease (increase) in the time to event) relative to the seizure rate in the reference group. The reference group contains individuals with partial seizures pre-randomisation, with a normal EEG and randomised to deferred treatment.

The parameter estimates for the times from first to second seizure, for the Log-Logistic and Lomax distributions are given in Table 3. The AIC for the Log-logistic distribution is 6024, and the AIC for the Lomax distribution is 5984, meaning that again the Lomax

distribution is the preferred of the two distributions.

Computing the Wald test statistics for each of the estimated regression coefficients presented in Table 3, we conclude that only θ_{t-c} is significant in both of the survival distributions. Additionally, $\theta_{t-c\times trt}$ is significant in the Log-logistic distribution only. This time, however, the lack of significant covariates is not as surprising, as the exploratory analysis suggested that most covariates that were statistically significant in determining the times to first seizure failed to be significant when considering the times from first to second seizure.

Recall that in the joint model each of the survival times were assumed to be Exponentially distributed, with the same rate. The results presented in Tables 2 and 3 are an initial indication that assuming the post-randomisation survival times to be identically distributed may be inaccurate. The exploratory analysis and initial modelling has suggested that there may be a change of seizure rate not only at randomisation, but also following the first seizure post-randomisation.

5.3 The Joint Model

Despite the exploratory analysis and initial modelling highlighting potential flaws in the joint model outlined in §3, we continue regardless and present results obtained through the fitting of this model. Two versions of the joint model were considered, the first jointly modelled the pre-randomisation event counts and times to first post-randomisation seizure and the second additionally included the times from first to second post-randomisation seizure. Estimated regression coefficients for the two fitted models are given in Table 4.

It is encouraging to note that the estimated regression coefficients contained in λ_i are very similar to those obtained through the Negative Binomial marginal count model, presented in Table 1. Wald tests on each of the regression coefficients in ψ_i show that in the joint model all covariates but the pre-randomisation seizure types are statistically significant. It may be surprising to conclude that the pre-randomisation seizure types are not statistically significant in the model, but note that their interactions with treatment policy and EEG outcome are highly statistically significant.

Comparing the statistically significant variables appearing in ψ_i for each of the two models considered, the estimated regression coefficients obtained through the joint modelling of the pre-randomisation event counts, and two post-randomisation survival times are closer to zero than those estimates obtained through the joint model that considers time to first seizure only. Recall that the explanatory analysis suggested that covariates that were statistically significant in determining the times to first seizure post-randomisation may not be significant when analysing the times from first to second post-randomisation seizures.

This may explain the averaging down effect observed here and provides further evidence that the assumption of a constant ψ_i post-randomisation may be flawed.

	Regression Coefficient	Estimates (standard errors) for the following models:	
		Time to First Seizure Only	Time to First and Second Seizure
	α	0.846 (0.037)	0.864 (0.037)
λ_i	$\beta_{1,0}$	-2.935 (0.141)	-3.143 (0.136)
	$\beta_{1,t-c}$	-0.629 (0.153)	-0.525 (0.148)
	$\beta_{1,2°t-c}$	-0.600 (0.159)	-0.530 (0.155)
	$\beta_{1,gen}$	0.693 (0.224)	0.828 (0.216)
	$\beta_{1,par}$	reference	reference
	$\beta_{1,oth}$	-0.200 (0.259)	0.006 (0.254)
ψ_i	$\beta_{2,0}$	-3.833 (0.378)	-2.922 (0.288)
	$\beta_{2,trt}$	1.827 (0.383)	1.192 (0.303)
	$\beta_{2,t-c}$	0.115 (0.389)	-0.461 (0.300)
	$\beta_{2,2°t-c}$	0.120 (0.401)	-0.132 (0.309)
	$\beta_{2,gen}$	-0.062 (0.598)	-0.106 (0.462)
	$\beta_{2,par}$	reference	reference
	$\beta_{2,oth}$	1.503 (0.620)	0.445 (0.504)
	$\beta_{2,eeg}$	-0.896 (0.392)	-1.232 (0.308)
	$\beta_{2,t-c\times trt}$	-2.012 (0.387)	-1.247 (0.306)
	$\beta_{2,2°t-c\times trt}$	-2.201 (0.400)	-1.615 (0.315)
	$\beta_{2,gen\times trt}$	-2.674 (0.562)	-2.209 (0.424)
	$\beta_{2,par\times trt}$	reference	reference
	$\beta_{2,oth\times trt}$	-3.454 (0.730)	-2.822 (0.605)
	$\beta_{2,eeg\times trt}$	-0.848 (0.199)	-0.464 (0.162)
	$\beta_{2,t-c\times eeg}$	1.610 (0.405)	1.643 (0.321)
	$\beta_{2,2°t-c\times eeg}$	2.395 (0.418)	2.091 (0.331)
	$\beta_{2,gen\times eeg}$	2.414 (0.625)	2.212 (0.483)
	$\beta_{2,par\times eeg}$	reference	reference
	$\beta_{2,oth\times eeg}$	1.894 (0.756)	2.674 (0.633)
$-$Log-likelihood (d.f.)		9669 (1398)	13353 (1398)

Table 4: Estimated regression coefficients for the joint models. The term λ_i contains parameter estimates corresponding to the effect of covariates on the underlying event rate and ψ_i contains parameter estimates corresponding to the effect of covariates on the post-randomisation reduction in event rates. A regression coefficient > 0 (< 0) would indicate an increased (decreased) seizure rate relative to the seizure rate in the reference group. The reference group contains individuals with partial seizures pre-randomisation, with a normal EEG and randomised to deferred treatment.

To gain a better understanding of the estimated regression coefficients, given in Table

4, subsequent estimates of the pre and post-randomisation seizure rates for the different seizure types, EEG outcomes and treatment policies can be calculated. We shall use the estimates given by the joint modelling of the pre-randomisation event counts and post-randomisation time to first seizure only.

Seizure Type	$\widehat{\lambda}_i$ (95% C.I.)
Tonic-Clonic	0.0283 (0.0249,0.0322)
2° Tonic-Clonic	0.0292 (0.0249,0.0341)
Generalised	0.1063 (0.0750,0.1505)
Partial	0.0531 (0.0401,0.0704)
Other	0.0435 (0.0281,0.0673)

Table 5: The values of $\widehat{\lambda}_i$ represent the expected pre-randomisation seizure rate per unit time.

Table 5 gives expected pre-randomisation seizure rates per unit time, for the different seizure types. Those individuals with generalised seizures pre-randomisation can typically expect to have the highest seizure rate, with those experiencing tonic-clonic seizures only and secondary tonic-clonic seizures having the lowest rate.

Seizure Type	$\widehat{\psi}_i$ (95% C.I.)	
	Abnormal EEG	
	Immediate	Deferred
Tonic-Clonic	0.018 (0.013,0.024)	0.050 (0.038,0.065)
2° Tonic-Clonic	0.032 (0.023,0.044)	0.109 (0.079,0.152)
Generalised	0.017 (0.008,0.035)	0.093 (0.050,0.172)
Partial	0.024 (0.012,0.046)	0.009 (0.006,0.014)
Other	0.022 (0.006,0.085)	0.264 (0.085,0.821)
	Normal EEG	
	Immediate	Deferred
Tonic-Clonic	0.020 (0.016,0.026)	0.024 (0.019,0.031)
2° Tonic-Clonic	0.017 (0.012,0.023)	0.024 (0.018,0.034)
Generalised	0.009 (0.003,0.024)	0.020 (0.008,0.053)
Partial	0.135 (0.066,0.274)	0.022 (0.010,0.046)
Other	0.019 (0.007,0.054)	0.097 (0.036,0.263)

Table 6: The values of $\widehat{\psi}_i$ represent the expected change in seizure rate post-randomisation.

Table 6 gives estimates of the expected change in seizure rate post-randomisation, stratified by seizure type, EEG outcome and treatment policy. As an example consider a person presenting generalised seizures pre-randomisation, with an abnormal EEG and randomised to deferred treatment. Table 5 tells us that their corresponding $\widehat{\lambda}_i$ is 0.1063, which equates to a seizure approximately every 9.4 days. Their subsequent $\widehat{\psi}_i$, from Table 6, is 0.093,

meaning that post-randomisation they should expect to have seizures about 9.3% as often as pre-randomisation. Recall that the post-randomisation seizure rate per unit time is given by $\widehat{\lambda}_i\widehat{\psi}_i = 0.1063 \times 0.093 = 0.01$, which equates to one seizure approximately every 100 days.

Looking at the values of $\widehat{\psi}_i$ presented in Table 6, treatment policy does not appear to be statistically significant for those individuals with a normal EEG. Additionally, those individuals having an abnormal EEG, but allocated to immediate treatment can expect to have a post-randomisation seizure rate in line with those presenting a normal EEG. For those with an abnormal EEG immediate treatment is favoured for all groups except partial, where no significant difference between treatment policies is observed. These conclusions are in line with the exploratory analysis.

Those with tonic-clonic seizures (either primary or secondary) have very similar pre-randomisation seizure rates, but for those presenting an abnormal EEG, the seizure rates post-randomisation are different. Finally, although those with generalised seizures pre-randomisation generally have the highest seizure rate, examination of Table 6 tells us that these individuals can generally expect to see the greatest reduction in seizure rates post-randomisation.

6 Extensions

Examination of the results obtained following the implementation of the joint model developed in §3 has highlighted possible limitations. In this next section we discuss a number of these limitations and explore possible solutions to the problems that they present.

6.1 Adjustments to the Pre-Randomisation Seizure Rates

It is important to note that 812 of the 1425 individuals included in the exploratory analysis presented only a single seizure pre-randomisation. The period of time from this single seizure to randomisation, for these individuals, ranged from the same day to 464 days, with the median number of days being 27. For the majority of those individuals with only one seizure pre-randomisation, their associated period of time from first seizure to randomisation may be inaccurately small, possibly representing how long it took for them to arrange an appointment with their GP. This results in imprecise estimates of their associated underlying seizure rates and an ensuing overestimation of the seizure rate reductions. Following discussions with clinicians, we subsequently made adjustments to the values of u_i in the data set so that $u_i \geq 182$. A sensitivity analysis shall be presented in §6.1.1.

After adjusting the data we reapplied the joint model (pre-randomisation event count

and time to first seizure post-randomisation only) and used the subsequent regression coefficients (which can be found in the Appendix, Table 9) to obtain updated estimates of $\widehat{\lambda}_i$ and $\widehat{\psi}_i$. The ensuing $\widehat{\lambda}_i$ can be found in Table 7.

Seizure Type	$\widehat{\lambda}_i$ (95% C.I.)
Tonic-Clonic	0.0056 (0.0051,0.0061)
2° Tonic-Clonic	0.0081 (0.0073,0.0090)
Generalised	0.0534 (0.0421,0.0678)
Partial	0.0161 (0.0134,0.0195)
Other	0.0220 (0.0163,0.0296)

Table 7: The values of $\widehat{\lambda}_i$ represent the expected pre-randomisation seizure rate per unit time, with a minimum pre-randomisation period of 182 days.

Adjustments to the pre-randomisation periods has resulted in significant reductions in the pre-randomisation seizure rates across all seizure type groups. Those with generalised seizures pre-randomisation can still expect to have the highest pre-randomisation seizure rate, but now the seizure rates for those with tonic-clonic seizures only and secondary tonic-clonic seizures are significantly different.

Seizure Type	$\widehat{\psi}_i$ (95% C.I.)	
	Abnormal EEG	
	Immediate	Deferred
Tonic-Clonic	0.079 (0.061,0.103)	0.189 (0.149,0.240)
2° Tonic-Clonic	0.104 (0.079,0.138)	0.284 (0.211,0.383)
Generalised	0.026 (0.014,0.050)	0.104 (0.058,0.187)
Partial	0.075 (0.043,0.129)	0.048 (0.031,0.075)
Other	0.057 (0.017,0.193)	0.279 (0.101,0.770)
	Normal EEG	
	Immediate	Deferred
Tonic-Clonic	0.087 (0.069,0.110)	0.113 (0.090,0.140)
2° Tonic-Clonic	0.057 (0.043,0.076)	0.084 (0.064,0.112)
Generalised	0.028 (0.013,0.062)	0.060 (0.026,0.138)
Partial	0.186 (0.103,0.339)	0.066 (0.034,0.127)
Other	0.059 (0.024,0.148)	0.157 (0.062,0.398)

Table 8: The values of $\widehat{\psi}_i$ represent the expected change in seizure rate post-randomisation, with a minimum pre-randomisation period of 182 days.

What is most apparent when looking at Table 8 is that the magnitude of the seizure rate reductions are not as large as those presented in Table 6. This is an immediate consequence of the reductions in pre-randomisation seizure rates observed in Table 7. Looking at the values of $\widehat{\psi}_i$ presented in Table 8 we can see that, as in Table 6, treatment policy does not appear to be significant for those individuals with a normal EEG. Additionally, those

having an abnormal EEG but allocated to immediate treatment have a post-randomisation seizure rate in line with those having a normal EEG. Once again, for those with an abnormal EEG immediate treatment is favoured for all groups except partial, where no significant difference between treatment policies is observed.

6.1.1 Sensitivity Analysis

As a sensitivity analysis to the choice of 182 days as the minimum period pre-randomisation, the data were re-analysed with $u_i \geq 91$ and $u_i \geq 365$. The resulting regression coefficients from these adjustments can be found in the Appendix, along with their associated $\widehat{\lambda}_i$ and $\widehat{\psi}_i$. The magnitudes of differences observed in seizure rates between the groups were maintained through each adjustment.

The log-likelihoods associated with each model would suggest that having a minimum pre-randomisation period of 365 days is optimal. Further inspection of the log-likelihoods however, suggests that the likelihood function is very flat, hence the decision was made to take the clinicians suggestion of a minimum pre-randomisation period of 182 days. All future analysis of the MESS data will be with this adjustment.

6.2 Time Varying Seizure Rate

The results have suggested that the iid assumption for Y_{1i} and Y_{2i} may not be accurate. Instead we consider a model that allows both Y_{1i} and Y_{2i} to be Exponentially distributed, and consider the following adjustment:

$$\begin{aligned} X_i \mid \nu_i &\sim \text{Poisson}(\lambda_i u_i \nu_i), \\ Y_{1i} \mid \nu_i &\sim \text{Exponential}(\lambda_i \psi_{1i} \nu_i), \\ Y_{2i} \mid \nu_i &\sim \text{Exponential}(\lambda_i \psi_{2i} \nu_i), \\ \nu_i &\sim \text{Gamma}(\alpha, \alpha), \end{aligned}$$

with $\psi_{1i} \neq \psi_{2i}$. The joint density of the survival times is the product of the densities of Y_{1i} and Y_{2i}, so that the joint model is specified by the following equations:

$$\begin{aligned} f_{X|\nu}(x_i \mid \nu_i; \lambda_i, u_i) &= \frac{(\lambda_i u_i \nu_i)^{x_i} \exp(-\lambda_i u_i \nu_i)}{x_i!}, \\ f_{Y_1,Y_2|\nu}(y_{1i}, y_{2i} \mid \nu_i; \lambda_i, \psi_i) &= (\lambda_i \nu_i)^2 \psi_{1i} \psi_{2i} \exp(-\lambda_i \nu_i (\psi_{1i} y_{1i} + \psi_{2i} y_{2i})), \\ g_\nu(\nu_i; \alpha) &= \frac{\alpha^\alpha \nu_i^{\alpha-1} \exp(-\alpha \nu_i)}{\Gamma(\alpha)}, \end{aligned}$$

where $\lambda_i = \exp(\boldsymbol{\beta}_1' \boldsymbol{z}_{1i})$, $\psi_{1i} = \exp(\boldsymbol{\beta}_2' \boldsymbol{z}_{2i})$ and $\psi_{2i} = \exp(\boldsymbol{\beta}_2' \boldsymbol{z}_{2i} + \ln(\rho)) = \rho \psi_{1i}$.

Integrating the random effect term out of the joint density of the survival times, Y_{1i},

Y_{2i} and ν_i, the unconditional joint distribution of the Y_{ji}, $j = 1, 2$ remains the bivariate Lomax distribution, however the density and survivor functions are now given by:

$$f_{Y_1,Y_2}(y_{1i}, y_{2i}; \lambda_i, \psi_i, \alpha) = \frac{\alpha+1}{\alpha}\lambda_i^2\psi_{1i}\psi_{2i}\left\{1 + \frac{\lambda_i\psi_{1i}y_{1i}}{\alpha} + \frac{\lambda_i\psi_{2i}y_{2i}}{\alpha}\right\}^{-(\alpha+2)}, \quad (12)$$

$$S_{Y_1,Y_2}(y_{1i}, y_{2i}; \lambda_i, \psi_i, \alpha) = \left\{1 + \frac{\lambda_i\psi_{1i}y_{1i}}{\alpha} + \frac{\lambda_i\psi_{2i}y_{2i}}{\alpha}\right\}^{-\alpha}. \quad (13)$$

Proceeding in the same manner as detailed in §3, considering the different censoring patterns separately, allows us to formulate the log-likelihood for the observed data $\mathcal{D}$ on all the n individuals, obtaining

$$\begin{aligned}\ell(\alpha, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \rho \mid \mathcal{D}) = \sum_{i=1}^{n}\Bigg\{&\left[\sum_{k=0}^{x_i-1}\ln(\alpha+k)\right] + (x_i + \delta_{1i}(1+\delta_{2i}))\ln(\lambda_i) + x_i\ln(u_i)\\ &- \ln(x_i!) + \alpha\ln(\alpha) + \delta_{1i}(1+\delta_{2i})\ln(\psi_{1i}) + \delta_{1i}\delta_{2i}\ln(\rho)\\ &+\delta_{1i}\ln(x_i+\alpha) + \delta_{1i}\delta_{2i}\ln(x_i+\alpha+1)\\ &-(x_i+\alpha+\delta_{1i}(1+\delta_{2i}))\ln\left(\lambda_i u_i + \lambda_i\psi_{1i}(y_{1i}+\rho y_{2i}) + \alpha\right)\Bigg\}. \quad (14)\end{aligned}$$

First and second derivatives of this function can be derived, allowing inference on the parameters α, $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ and ρ, using a numerical method such as Newton-Raphson.

The subsequent regression coefficients for this joint model can be found in Table 16 in the Appendix. A Wald test on the estimated ρ confirms that this variable is highly statistically significant, with the seizure rate following the first post-randomisation seizure likely to be twice as high as the seizure rate immediately succeeding randomisation. The derivation of this adjusted model allows us to conduct a hypothesis test, comparing the model presented in Table 4 with the model developed here. The corresponding deviance is given by $-2\{-13353 + 12119\} = 2468$, which provides overwhelming support for the introduction of the parameter ρ.

We do not believe that this model is now sufficient for modelling the data, we merely use this example to illustrate that the identically distributed assumption for the two survival times is flawed. We aim to develop a model that has $\psi_{1i} = \exp(\boldsymbol{\beta}_2'\boldsymbol{z}_{2i})$ and $\psi_{2i} = \exp(\boldsymbol{\beta}_3'\boldsymbol{z}_{3i})$, allowing different covariates to appear in ψ_{1i} and ψ_{2i}.

6.3 Cure Rate Models

An additional potential extension to the statistical model used to analyse the data, is the inclusion of possible cure rates. The magnitude of the reductions in seizure rates post-randomisation suggests that there may be a substantial proportion of the population that we should regard as cured. Berg and Shinnar (1991) noted that, on average, around 50%

of people do not experience seizure recurrence after a single, untreated seizure. Recall that over half of the 1425 individuals for which exploratory analysis was carried out presented only a single seizure pre-randomisation. It is therefore not unreasonable to suspect that a substantial proportion of the sample would never have a seizure post-randomisation, regardless of the length of time for which they were followed. Alternatively Maller and Zhou (1996) encourages us to think of these individuals as cured, or immune to seizure recurrence. If survival data does indeed have a proportion that are immune to the event of interest, considering a model that ignores this may give misleading results. More specifically, ignoring any potential cure fraction could result in underestimates of the post-randomisation seizure rates, thus contributing to the magnitude of seizure rate reductions that have been observed.

A proper survival distribution should have total mass 1 with the resulting Kaplan-Meier curve having its asymptote at zero. A survival distribution that is improper allows, formally, infinite survival times. Cure rate models allow the quantity $p = F(\infty) = \lim_{t\to\infty} F(t)$ (where $F(t)$ is the cumulative distribution function of the survival times) to be strictly less than 1, corresponding to the presence of immunes in the population.

Suppose t_i^* is the true survival time for individual i, which is only observed if it does not exceed individual i's censoring time c_i, otherwise we observe c_i. Consequently, the actual, observed survival time for individual i can be expressed as $t_i = \min(t_i^*, c_i)$. To formulate the probabilistic mechanism that allows the true survival times t_i^* to be infinite first assume that individual i has an associated Bernoulli random variable, B_i, taking the value 1 if individual i is susceptible to the event of interest, and with $B_i = 0$ corresponding to an immune individual. Additionally, $p < 1$ represents the proportion of susceptibles in the population, so that

$$B_i = \begin{cases} 1 & \text{with probability} \quad p, \\ 0 & \text{with probability} \quad 1-p. \end{cases}$$

In reality we do not know whether an individual is immune or not, so B_i is not observed. Susceptible individuals are assumed to have a proper cumulative survival distribution $F_0(t)$, with $F_0(\infty) = 1$. Formally, individuals with $B_i = 0$ have $t_i^* = \infty$, hence, for all $t \geq 0$

$$\begin{aligned} \mathbb{P}\{t_i^* \leq t \mid B_i = 1\} &= F_0(t), \\ \mathbb{P}\{t_i^* \leq t \mid B_i = 0\} &= 0, \end{aligned}$$

These probabilities imply that, for all $t \geq 0$, the cumulative distribution function of the

true survival times t_i^* is

$$\begin{aligned} F(t) = \mathbb{P}\{t_i^* \leq t\} &= \mathbb{P}\{t_i^* \leq t \mid B_i = 1\}\mathbb{P}\{B_i = 1\} + \mathbb{P}\{t_i^* \leq t \mid B_i = 0\}\mathbb{P}\{B_i = 0\} \\ &= pF_0(t) + 0 \\ &= pF_0(t). \end{aligned}$$

Consequently, for all $t \geq 0$

$$F_0(t) = \frac{F(t)}{p} = \frac{F(t)}{F(\infty)}.$$

To ensure that p remains within the interval $[0, 1]$ the following reparameterisation is often considered:

$$\kappa = \ln\left(\frac{p}{1-p}\right).$$

We have established that the decision as to whether or not immunes are present is a critical one in the analysis of survival data. However, an important factor in the determination of an immune component is the idea of sufficient follow-up: has there been sufficient follow-up in the sample so that we can detect, with confidence, immunes in the population. Maller and Zhou (1996) discusses how to formally test for the presence of immunes and sufficient follow-up, although details are excluded here.

Both the Kaplan-Meier curves in Figure 7 seem to level off well above zero, suggesting that there may be an immune component present for both time to first seizure post-randomisation and time from first to second seizure.

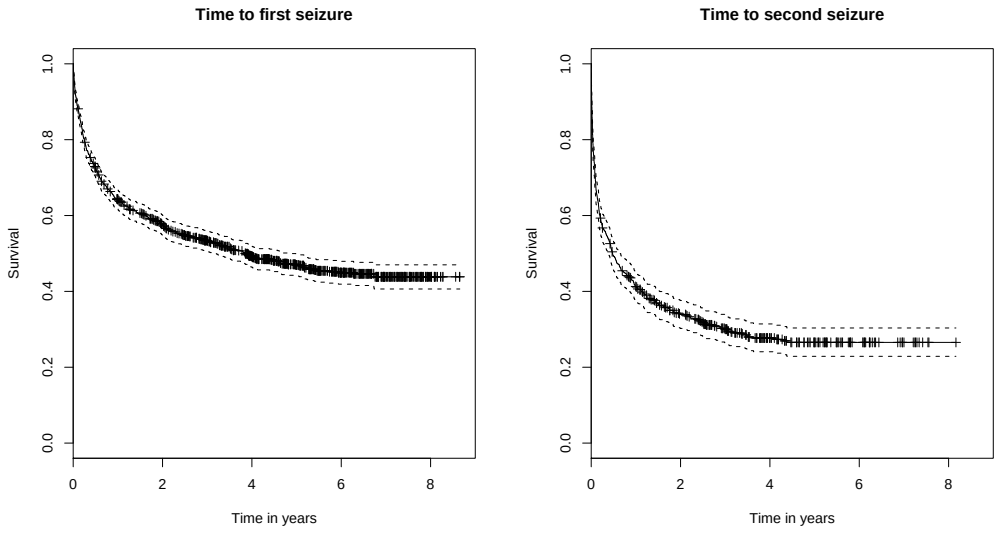

Figure 7: Kaplan-Meier curves for time to first and second seizures (with 95% CI). Curves are marked at each censoring time which is not also a death time

Standard software (Peng) allows the fitting of various parametric mixture models, including the Log-logistic, for the estimation of cure rates. We therefore proceed to fit a Log-logistic mixture model that allows for the existence of a cure fraction. This model was considered for both time to first seizure post-randomisation and time from first to second seizure, the subsequent regression coefficients are presented in Table 17, which can be found in the Appendix.

Wald tests on the estimates for κ in each of the two models confirm that they are highly statistically significant, implying that there is a cure fraction present for both the survival times. The subsequent cure fractions can be calculated by considering the following one-to-one transformation:

$$\begin{aligned} 1-p &= 1-\frac{\exp(\kappa)}{1+\exp(\kappa)} \\ &= \frac{1}{1+\exp(\kappa)} \end{aligned}$$

The estimated cure fractions are 21.7% for time to first seizure, and 18.8% for time from first to second seizure. These values are slightly lower than suggested by the Kaplan-Meier curves in Figure 7, but still provide sufficient evidence to suggest that we should incorporate a cure fraction into the joint model.

6.4 One-Inflated, Zero-Truncated Poisson Distribution

It has already been noted that over half of the participants randomised presented only a single seizure pre-randomisation. This excess of ones that the data displays is not accounted for in the model developed in Section §3. Additionally, recall that the eligibility criterion for the MESS trial specified that participants should have had at least one epileptic seizure pre-randomisation. A one-inflated, zero-truncated Poisson distribution is considered for the pre-randomisation event counts.

Recall that the density for a standard Poisson(λ) distribution is given by:

$$\mathrm{Po}(x;\lambda) = \frac{\lambda^x \exp(-\lambda)}{x!}.$$

The zero-truncated Poisson is a model for count data that is truncated at zero. By the definition of conditional probability, the probability function of the zero-truncated Poisson(λ) random variable satisfies

$$\mathbb{P}(X=k) = \mathbb{P}(W=k \mid W>0) = \frac{\mathbb{P}(W=k)}{1-\mathbb{P}(W=0)},$$

where $W \sim$Poisson(λ). The resulting density function for the zero-truncated Poisson(λ) distribution is

$$\text{ZTP}(x;\lambda) = \frac{\lambda^x \exp(-\lambda)}{x!(1-\exp(-\lambda))} = \frac{\lambda^x}{x!(\exp(\lambda)-1)}.$$

The one-inflated, zero-truncated Poisson distribution is a model for data that exhibits excess ones and is truncated at zero. The model assumes that, with probability π, the only possible observation is 1, and with probability $1-\pi$ a zero-truncated Poisson(λ) random variable is observed. Hence,

$$\begin{aligned} X = 1 \quad &\text{with probability} \quad \pi + (1-\pi)\frac{\lambda}{(\exp(\lambda)-1)}, \\ X = k > 1 \quad &\text{with probability} \quad (1-\pi)\frac{\lambda^k}{k!(\exp(\lambda)-1)}, \end{aligned}$$

giving, for $x > 1$,

$$f_X(x;\lambda,\pi) = \pi \mathbb{I}_{[x=1]} + (1-\pi)\text{ZTP}(x;\lambda), \tag{15}$$

where $\mathbb{I}_{[x=1]}$ is the indicator function taking the value 1 when $x = 1$ and zero otherwise.

7 Discussion and Conclusions

Examination of the results presented in §5 provides sufficient evidence for the use of the joint model developed in §3, over standard survival techniques. The inclusion of additional information in the joint model resulted in an increase in power, which consequently meant that statistically significant covariate effects, not recognised by the standard survival distributions, could be affirmed. Despite this support for the joint model, further examination of the results obtained highlighted interesting characteristics within the data are not present in the model.

A substantial number of individuals entered the MESS trial having had only a single seizure, which meant that there were associated difficulties in estimating pre-randomisation seizure rates. This issue was resolved in §6.1 by adjusting the minimum period of time pre-randomisation, over which seizures were observed, to 182 days. The improvements to the subsequent $\widehat{\lambda}_i$ and $\widehat{\psi}_i$ can be seen in Tables 7 and 8. A further concern discussed was the fact that the model did not account for the excess of single seizures pre-randomisation displayed by the data. This was addressed in §6.4 by formulating a one-inflated, zero-truncated Poisson distribution, but the consequences of this distribution for the pre-randomisation event count data has yet to be investigated.

It was established in §6.3 and Table 17 that there is sufficient evidence to support the presence of a cure fraction in the population, with a proportion of individuals likely to be

'immune' from future seizures post-randomisation. Initial analysis has confirmed that the incorporation of a cure rate would be justifiable.

Finally, it was immediately obvious from the results discussed in §5 that the assumption of identically distributed post-randomisation survival times was not supported. In §6.2 and Table 16 it was formally proved that there was overwhelming evidence to support the inclusion of an additional parameter, ρ, where $\psi_{1i} = \exp(\boldsymbol{\beta}_2' \boldsymbol{z}_{2i})$ and $\psi_{2i} = \rho\psi_{1i}$. We wish to proceed with the development of our model by first considering an alternative form for ψ_{2i}, allowing this parameter to depend on covariates that are separate from those appearing in ψ_{1i}. More specifically, we shall specify the joint model by the following equations:

$$\begin{aligned} f_{X|\nu}(x_i \mid \nu_i; \lambda_i, u_i) &= \frac{(\lambda_i u_i \nu_i)^{x_i} \exp(-\lambda_i u_i \nu_i)}{x_i!}, \\ f_{Y_1, Y_2|\nu}(y_{1i}, y_{2i} \mid \nu_i; \lambda_i, \psi_i) &= (\lambda_i \nu_i)^2 \psi_{1i} \psi_{2i} \exp(-\lambda_i \nu_i (\psi_{1i} y_{1i} + \psi_{2i} y_{2i})), \\ g_\nu(\nu_i; \alpha) &= \frac{\alpha^\alpha \nu_i^{\alpha-1} \exp(-\alpha\nu_i)}{\Gamma(\alpha)}, \end{aligned}$$

where $\lambda_i = \exp(\boldsymbol{\beta}_1' \boldsymbol{z}_{1i})$, $\psi_{1i} = \exp(\boldsymbol{\beta}_2' \boldsymbol{z}_{2i})$ and $\psi_{2i} = \exp(\boldsymbol{\beta}_3' \boldsymbol{z}_{3i})$.

Following this adjustment to the joint model we shall consider the inclusion of cure rates, before pursuing the one-inflated, zero-truncated Poisson distribution. To formally examine the performance of the models that have been developed to date, and models that continue to be explored we shall work to obtain residuals as a method of model checking.

Appendix

Adjustments to the Pre-Randomisation Seizure Rates

Minimum pre-randomisation period of 182 days

	Regression Coefficient	Estimates (standard errors) for the following models:	
		Negative Binomial GLM	Joint Model
	α	1.660 (0.078)	1.563 (0.074)
λ_i	$\beta_{1,0}$	-4.112 (0.091)	-4.127 (0.094)
	$\beta_{1,t-c}$	-1.073 (0.100)	-1.063 (0.104)
	$\beta_{1,2^\circ t-c}$	-0.679 (0.103)	-0.688 (0.107)
	$\beta_{1,gen}$	1.185 (0.148)	1.198 (0.152)
	$\beta_{1,par}$	reference	reference
	$\beta_{1,oth}$	0.301 (0.172)	0.310 (0.176)
ψ_i	$\beta_{2,0}$	-	-2.722 (0.330)
	$\beta_{2,trt}$	-	1.043 (0.333)
	$\beta_{2,t-c}$	-	0.539 (0.341)
	$\beta_{2,2^\circ t-c}$	-	0.249 (0.351)
	$\beta_{2,gen}$	-	-0.090 (0.522)
	$\beta_{2,par}$	-	reference
	$\beta_{2,oth}$	-	0.870 (0.567)
	$\beta_{2,eeg}$	-	-0.306 (0.342)
	$\beta_{2,t-c\times trt}$	-	-1.302 (0.337)
	$\beta_{2,2^\circ t-c\times trt}$	-	-1.434 (0.350)
	$\beta_{2,gen\times trt}$	-	-1.806 (0.500)
	$\beta_{2,par\times trt}$	-	reference
	$\beta_{2,oth\times trt}$	-	-2.016 (0.658)
	$\beta_{2,eeg\times trt}$	-	-0.612 (0.180)
	$\beta_{2,t-c\times eeg}$	-	0.824 (0.352)
	$\beta_{2,2^\circ t-c\times eeg}$	-	1.521 (0.364)
	$\beta_{2,gen\times eeg}$	-	0.859 (0.535)
	$\beta_{2,par\times eeg}$	-	reference
	$\beta_{2,oth\times eeg}$	-	0.881 (0.680)
$-$Log-likelihood (d.f.)		2833 (1414)	8580 (1398)

Table 9: Estimated regression coefficients for the joint models, with a minimum pre-randomisation period of 182 days. The term λ_i contains parameter estimates corresponding to the effect of covariates on the underlying event rate and ψ_i contains parameter estimates corresponding to the effect of covariates on the post-randomisation reduction in event rates. A regression coefficient > 0 (< 0) would indicate an increased (decreased) seizure rate relative to the seizure rate in the reference group. The reference group contains individuals with partial seizures pre-randomisation, with a normal EEG and randomised to deferred treatment.

Minimum pre-randomisation period of 91 days

	Regression Coefficient	Estimates (standard errors) for the following models:	
		Negative Binomial GLM	Joint Model
	α	1.443 (0.064)	1.394 (0.063)
λ_i	$\beta_{1,0}$	-3.744 (0.097)	-3.789 (0.101)
	$\beta_{1,t-c}$	-0.919 (0.106)	-0.898 (0.111)
	$\beta_{1,2^\circ t-c}$	-0.631 (0.110)	-0.633 (0.114)
	$\beta_{1,gen}$	1.080 (0.159)	1.121 (0.163)
	$\beta_{1,par}$	reference	reference
	$\beta_{1,oth}$	0.158 (0.186)	0.181 (0.188)
ψ_i	$\beta_{2,0}$	-	-3.164 (0.342)
	$\beta_{2,trt}$	-	1.223 (0.343)
	$\beta_{2,t-c}$	-	0.497 (0.352)
	$\beta_{2,2^\circ t-c}$	-	0.299 (0.362)
	$\beta_{2,gen}$	-	-0.100 (0.534)
	$\beta_{2,par}$	-	reference
	$\beta_{2,oth}$	-	1.024 (0.583)
	$\beta_{2,eeg}$	-	-0.343 (0.352)
	$\beta_{2,t-c\times trt}$	-	-1.476 (0.347)
	$\beta_{2,2^\circ t-c\times trt}$	-	-1.638 (0.359)
	$\beta_{2,gen\times trt}$	-	-2.005 (0.511)
	$\beta_{2,par\times trt}$	-	reference
	$\beta_{2,oth\times trt}$	-	-2.341 (0.675)
	$\beta_{2,eeg\times trt}$	-	-0.654 (0.183)
	$\beta_{2,t-c\times eeg}$	-	0.896 (0.3632)
	$\beta_{2,2^\circ t-c\times eeg}$	-	1.604 (0.375)
	$\beta_{2,gen\times eeg}$	-	1.053 (0.552)
	$\beta_{2,par\times eeg}$	-	reference
	$\beta_{2,oth\times eeg}$	-	1.066 (0.700)
$-$Log-likelihood (d.f.)		2909 (1414)	8689 (1398)

Table 10: Estimated regression coefficients for the joint models, with a minimum pre-randomisation period of 91 days. The term λ_i contains parameter estimates corresponding to the effect of covariates on the underlying event rate and ψ_i contains parameter estimates corresponding to the effect of covariates on the post-randomisation reduction in event rates. A regression coefficient > 0 (< 0) would indicate an increased (decreased) seizure rate relative to the seizure rate in the reference group. The reference group contains individuals with partial seizures pre-randomisation, with a normal EEG and randomised to deferred treatment.

Seizure Type	$\widehat{\lambda}_i$ (95% C.I.)
Tonic-Clonic	0.0092 (0.0084,0.0101)
2° Tonic-Clonic	0.0120 (0.0108,0.0134)
Generalised	0.0694 (0.0537,0.0896)
Partial	0.0226 (0.0185,0.0277)
Other	0.0271 (0.0197,0.0372)

Table 11: The values of $\widehat{\lambda}_i$ represent the expected pre-randomisation seizure rate per unit time, with a minimum pre-randomisation period of 91 days.

Seizure Type	$\widehat{\psi}_i$ (95% C.I.)	
	Abnormal EEG	
	Immediate	Deferred
Tonic-Clonic	0.049 (0.038,0.063)	0.121 (0.095,0.154)
2° Tonic-Clonic	0.069 (0.052,0.092)	0.201 (0.148,0.273)
Generalised	0.023 (0.012,0.043)	0.095 (0.052,0.173)
Partial	0.053 (0.030,0.094)	0.030 (0.019,0.046)
Other	0.041 (0.012,0.142)	0.242 (0.084,0.698)
	Normal EEG	
	Immediate	Deferred
Tonic-Clonic	0.054 (0.043,0.068)	0.070 (0.055,0.087)
2° Tonic-Clonic	0.038 (0.028,0.051)	0.057 (0.043,0.076)
Generalised	0.021 (0.009,0.049)	0.047 (0.020,0.109)
Partial	0.144 (0.077,0.268)	0.042 (0.021,0.084)
Other	0.039 (0.015,0.098)	0.118 (0.045,0.306)

Table 12: The values of $\widehat{\psi}_i$ represent the expected change in seizure rate post-randomisation, with a minimum pre-randomisation period of 91 days.

Minimum pre-randomisation period of 365 days

	Regression Coefficient	Estimates (standard errors) for the following models:	
		Negative Binomial GLM	Joint Model
	α	1.797 (0.088)	1.672 (0.081)
λ_i	$\beta_{1,0}$	-4.586 (0.087)	-4.593 (0.090)
	$\beta_{1,t-c}$	-1.163 (0.096)	-1.151 (0.099)
	$\beta_{1,2^\circ t-c}$	-0.685 (0.099)	-0.687 (0.102)
	$\beta_{1,gen}$	1.453 (0.141)	1.461 (0.146)
	$\beta_{1,par}$	reference	reference
	$\beta_{1,oth}$	0.393 (0.164)	0.402 (0.170)
ψ_i	$\beta_{2,0}$	-	-2.192 (0.332)
	$\beta_{2,trt}$	-	0.898 (0.327)
	$\beta_{2,t-c}$	-	0.541 (0.332)
	$\beta_{2,2^\circ t-c}$	-	0.1759 (0.342)
	$\beta_{2,gen}$	-	-0.386 (0.518)
	$\beta_{2,par}$	-	reference
	$\beta_{2,oth}$	-	0.778 (0.559)
	$\beta_{2,eeg}$	-	-0.288 (0.334)
	$\beta_{2,t-c\times trt}$	-	-1.151 (0.331)
	$\beta_{2,2^\circ t-c\times trt}$	-	-1.256 (0.344)
	$\beta_{2,gen\times trt}$	-	-1.665 (0.493)
	$\beta_{2,par\times trt}$	-	reference
	$\beta_{2,oth\times trt}$	-	-1.809 (0.650)
	$\beta_{2,eeg\times trt}$	-	-0.602 (0.178)
	$\beta_{2,t-c\times eeg}$	-	0.792 (0.344)
	$\beta_{2,2^\circ t-c\times eeg}$	-	1.490 (0.357)
	$\beta_{2,gen\times eeg}$	-	0.781 (0.526)
	$\beta_{2,par\times eeg}$	-	reference
	$\beta_{2,oth\times eeg}$	-	0.706 (0.667)
$-$Log-likelihood (d.f.)		2799 (1414)	8530 (1398)

Table 13: Estimated regression coefficients for the joint models, with a minimum pre-randomisation period of 365 days. The term λ_i contains parameter estimates corresponding to the effect of covariates on the underlying event rate and ψ_i contains parameter estimates corresponding to the effect of covariates on the post-randomisation reduction in event rates. A regression coefficient > 0 (< 0) would indicate an increased (decreased) seizure rate relative to the seizure rate in the reference group. The reference group contains individuals with partial seizures pre-randomisation, with a normal EEG and randomised to deferred treatment.

Seizure Type	$\widehat{\lambda}_i$ (95% C.I.)
Tonic-Clonic	0.0032 (0.0029,0.0035)
2° Tonic-Clonic	0.0051 (0.0046,0.0056)
Generalised	0.0436 (0.0347,0.0549)
Partial	0.0101 (0.0085,0.0121)
Other	0.0151 (0.0113,0.0202)

Table 14: The values of $\widehat{\lambda}_i$ represent the expected pre-randomisation seizure rate per unit time, with a minimum pre-randomisation period of 365 days.

Seizure Type	$\widehat{\psi}_i$ (95% C.I.)	
	Abnormal EEG	
	Immediate	Deferred
Tonic-Clonic	0.135 (0.104,0.175)	0.318 (0.252,0.402)
2° Tonic-Clonic	0.170 (0.129,0.224)	0.443 (0.329,0.596)
Generalised	0.032 (0.016,0.061)	0.124 (0.072,0.217)
Partial	0.113 (0.066,0.193)	0.084 (0.054,0.129)
Other	0.081 (0.024,0.273)	0.370 (0.140,0.979)
	Normal EEG	
	Immediate	Deferred
Tonic-Clonic	0.149 (0.119,0.187)	0.192 (0.155,0.239)
2° Tonic-Clonic	0.093 (0.070,0.124)	0.133 (0.101,0.176)
Generalised	0.035 (0.017,0.075)	0.076 (0.033,0.175)
Partial	0.274 (0.153,0.492)	0.112 (0.059,0.213)
Other	0.098 (0.040,0.241)	0.243 (0.096,0.613)

Table 15: The values of $\widehat{\psi}_i$ represent the expected change in seizure rate post-randomisation, with a minimum pre-randomisation period of 365 days.

Time Varying Seizure Rate

	Regression Coefficient	Estimates (standard errors)
	α	1.493 (0.070)
	ρ	2.203 (0.156)
λ_i	$\beta_{1,0}$	-4.142 (0.096)
	$\beta_{1,t-c}$	-1.047 (0.105)
	$\beta_{1,2^\circ t-c}$	-0.693 (0.109)
	$\beta_{1,gen}$	1.213 (0.155)
	$\beta_{1,par}$	reference
	$\beta_{1,oth}$	0.320 (0.180)
ψ_{1i}	$\beta_{2,0}$	-2.164 (0.255)
	$\beta_{2,trt}$	0.167 (0.267)
	$\beta_{2,t-c}$	-0.087 (0.265)
	$\beta_{2,2^\circ t-c}$	0.028 (0.272)
	$\beta_{2,gen}$	-0.062 (0.407)
	$\beta_{2,par}$	reference
	$\beta_{2,oth}$	-0.174 (0.475)
	$\beta_{2,eeg}$	-0.648 (0.271)
	$\beta_{2,t-c\times trt}$	-0.205 (0.270)
	$\beta_{2,2^\circ t-c\times trt}$	-0.638 (0.279)
	$\beta_{2,gen\times trt}$	-1.358 (0.379)
	$\beta_{2,par\times trt}$	reference
	$\beta_{2,oth\times trt}$	-1.350 (0.561)
	$\beta_{2,eeg\times trt}$	-0.181 (0.146)
	$\beta_{2,t-c\times eeg}$	0.878 (0.281)
	$\beta_{2,2^\circ t-c\times eeg}$	1.117 (0.291)
	$\beta_{2,gen\times eeg}$	0.716 (0.420)
	$\beta_{2,par\times eeg}$	reference
	$\beta_{2,oth\times eeg}$	1.732 (0.588)
$-$Log-likelihood (d.f.)		12119 (1397)

Table 16: Estimated regression coefficients for the joint model with a varying post-randomisation seizure rate. The term λ_i contains parameter estimates corresponding to the effect of covariates on the underlying event rate and ψ_{1i} contains parameter estimates corresponding to the effect of covariates on the post-randomisation reduction in event rates. Furthermore, $\psi_{2i} = \rho\psi_{1i}$. A regression coefficient > 0 (< 0) would indicate an increased (decreased) seizure rate relative to the seizure rate in the reference group. The reference group contains individuals with partial seizures pre-randomisation, with a normal EEG and randomised to deferred treatment.

Regression Coefficient	Estimates (standard errors) for the following models:	
	Time to First Seizure Only	Time from First to Second Seizure
θ_0	-0.514 (0.867)	-0.714 (1.035)
θ_{trt}	0.131 (1.246)	-1.050 (1.479)
θ_{t-c}	-0.668 (0.606)	-2.400 (0.826)
$\theta_{2^\circ t-c}$	-0.921 (0.623)	-1.199 (0.831)
θ_{gen}	-0.099 (1.006)	-0.764 (1.259)
θ_{par}	reference	reference
θ_{oth}	-0.020 (0.977)	-1.212 (1.678)
θ_{eeg}	0.391 (0.645)	-1.008 (0.834)
$\theta_{\ln(rate)}$	1.008 (0.153)	0.373 (0.172)
$\theta_{t-c\times trt}$	-0.669 (0.666)	1.223 (0.890)
$\theta_{2^\circ t-c\times trt}$	-0.689 (0.682)	0.611 (0.888)
$\theta_{gen\times trt}$	-0.806 (1.112)	-0.894 (1.311)
$\theta_{par\times trt}$	reference	reference
$\theta_{oth\times trt}$	-1.343 (1.204)	-2.493 (2.048)
$\theta_{eeg\times trt}$	-1.113 (0.350)	0.706 (0.470)
$\theta_{\ln(rate)\times trt}$	-0.045 (0.229)	0.023 (0.281)
$\theta_{t-c\times eeg}$	0.325 (0.674)	0.701 (0.869)
$\theta_{2^\circ t-c\times eeg}$	1.271 (0.706)	0.385 (0.889)
$\theta_{gen\times eeg}$	0.633 (1.142)	0.325 (1.481)
$\theta_{par\times eeg}$	reference	reference
$\theta_{oth\times eeg}$	0.329 (1.256)	2.316 (2.109)
κ	1.283 (0.243)	1.460 (0.187)
γ	0.725 (0.035)	0.763 (0.041)
$-$Log-likelihood (d.f.)	2177 (1400)	1255 (1400)

Table 17: Estimated regression coefficients for log-logistic mixture model fitted to the two survival times. The term κ is related to the cure fraction in the population. A regression coefficient, θ_k, > 0 (< 0) would indicate an increase (decrease) in the acceleration factor (hence a decrease (increase) in the time to event) relative to the seizure rate in the reference group. The reference group contains individuals with partial seizures pre-randomisation, with a normal EEG and randomised to deferred treatment.

References

Berg, A. T. and S. Shinnar (1991). The risk of seizure recurrence following a first unprovoked seizure: A quantitative review. *Neurology 41*(7), 965–972.

Chandra, B. (1992). First seizure in adults: to treat or not to treat. *Clinical Neurology and Neurosurgery 94(Suppl.)*, S61–S63.

Cowling, B. J., J. L. Hutton, and J. E. H. Shaw (2006). Joint modelling of event counts and survival times. *JRRSC Appl. Statist. 55*(1), 31–39.

Kim, L., A. Johnson, A. Marson, and D. Chadwick (2006). Prediction of risk of seizure recurrence after a single seizure and early epilepsy: further results from the mess trial. *The Lancet Neurology 5*, 317–322.

Maller, R. and X. Zhou (1996). *Survival Analysis with Long-Term Survivors*. Wiley series in Probability and Statistics. John Wiley and Sons.

Marson, A., A. Jacoby, A. Johnson, L. Kim, C. Gamble, and D. Chadwick (2005). Immediate versus deferred antiepileptic drug treatment for early epilepsy and single seizures: a randomised control trial. *The Lancet 365*, 2007–2013.

Nayak, T. K. (1987). Multivariate lomax distribution: properties and usefulness in reliability theory. *Journal of Applied Probability 24*, 170–177.

Peng, Y. gfcure. http://post.queensu.ca/∼pengp/software.html. Downloaded 14 July 2009.

Sahamotoa, Y., M. Ishiguro, and G. Kitagawa (1986). *Akaike Information Criterion Statistics*. Wiley series in Probability and Statistics. KFT Scientific.

Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society 54*(3), 426–482.